

No correlation between the Transit-Depth Metallicity of *Kepler* gas giant confirmed and candidates planets: A Bayesian Approach

Cyrine Nehmé^{1,2} and Paula Sarkis³

¹Department of Physics & Astronomy, Notre Dame University- Louaize, PO Box 72, Zouk Mikael, Lebanon

²CEA/DRF/Irfu/SAP, F-91191 Gif-sur-Yvette, France

³Universität Bern Space Research & Planetary Sciences Division Bern, Switzerland

E-mail: cnehme@ndu.edu.lb

Abstract. Previous attempts to study the correlation between the transit depth and the stellar metallicity of *Kepler* gas giant planets has led to different results. A weakly significant negative correlation was reported from the *Kepler*'s (Q1-Q12) gas giant candidates with estimated radii of $5\text{-}20R_{\odot}$ and $[\text{Fe}/\text{H}]$ values taken from the Kepler Input Catalog (KIC). With the release of the last *Kepler* catalog (Q1-Q17), we now have the largest homogeneous sample of exoplanets. This enables a solid statistical analysis of this correlation. In the present work, we revise this correlation, within a Bayesian framework, for two large homogeneous samples: confirmed and complete. We expand a Hierarchical model to account for false positives in the studied samples. Our statistical analysis reveals no correlation between the transit depth and the stellar metallicity. The fact that we found no evidence of such correlation will have implications for planet formation theory and interior structure of giant planets.

1. Introduction

NASA's Kepler Mission revolutionized the field of extra solar planets and now more than ever, it is possible to put statistical constraints on the observed planet properties and on theories of planet formation. Clues on the nature of giant planet formation might be revealed from the two correlations with stellar metallicity of main sequence stars hosting these planets. The first one is the correlation of the frequency of giant planets with stellar metallicity revealed by radial velocity surveys ([1]) and by transit surveys ([2]). The second correlation shows a positive trend between the mass of heavy-elements in giant planets and the stellar metallicity ([3]). This explosion of new information demonstrates the need to understand planet formation in general and presents an opportunity to compare observed trends to the theories of planet formation and evolution. Earlier attempts to study the correlation between the transit depth and the stellar metallicity of *Kepler*'s gas giant candidates has led to different results. For instance, [4] reported a negative correlation with a weak significant value. The author studied the transit depth of 218 giant planets from [5], (Q1-Q12) catalog with estimated radii of $5\text{-}20R_{\oplus}$ and the values of $[\text{Fe}/\text{H}]$ taken from the *Kepler* Input Catalog (KIC). [4] interpreted the negative correlation as evidence that metal rich planets of a given mass are denser than their metal poor counterparts leading to small radii ([6]). Here, we will use the latest available catalog Q1-Q17 ([7]) to study

correlation between the transit depth and stellar metallicity. Noting that a sample of stars with transiting planets may not accurately represent the true intrinsic distribution of the discovered planets. [8] reported the importance of including these effect since they can lead to biases in the properties of transiting planets and their host stars. For these reasons, we study only a subset of the target stars and the detected giant planets. We develop a flexible framework to account for uncertainties by expanding the Hierarchical Bayesian Model introduced by [9].

2. Selection criteria and complete sub-sample

We use the cumulative catalog of planets detected by the NASA *Kepler* mission which, as of April 2015, consisted of the latest Q1-Q17 catalog ([7]). Following [4] and [10] we define gas giant planets as planets having a radius between $5\text{-}20R_{\oplus}$. The stellar parameters were taken from the *Kepler* stellar Q1-Q16 database ([11]). We ended up by having 84 planets confirmed and 305 candidates.

With the goal of performing a robust statistical method, we prepared two different samples. The first sample consists of all the 84 confirmed giant planets within the latest catalog. The second sample contains a complete subsample of both confirmed and candidates. Hence, we performed cuts needed to take into account: the incompleteness of the catalog ([12]), the selection effects for host stars and planets candidates, the detection efficiency and the false positives. After performing all the cuts (table 1), we retain 105 confirmed + candidates planets. We believe that thi complete subsample better represents the true intrinsic distribution of *Kepler's* giant planets.

Table 1. Summary of the cuts performed to obtain a complete subsample

Parameter	Value
Stellar effective temperature, T_{eff}	4000 - 7000 K
Stellar gravity, $\log g$ (cm/s^{-2})	4.0 - 5.0
Stellar Radius, R_{\star}	0.7 - 1.4 R_{\odot}
Planetary Radius, R_p	5 - 20 R_{\oplus}
Orbital Period, P	< 90 days
Detection Efficiency, SNR	> 18.8
Kepler magnitude, K_p	< 16 mag

3. Method

Hierarchical Bayesian Modeling (hereafter HBM) allows for intrinsic scatter and heteroscedastic measurement errors i.e the uncertainties for each data points are different. We followed an approach similar to that proposed by [9]. We constructed the likelihood function in a simple way in order to relate the parameters of interest to the observed data, taking into account the measurements uncertainties. We extended the model to account for false positive by [13]. This updated model the HBM is used to study the correlation between the transit depth (δ) and the metallicity (FeH) of host stars. A graphical illustration of our HBM is given in Figure 1. Markov chain Monte Carlo (hereafter MCMC) was performed using the Python package PySTAN, a package for Bayesian inference. We ran models with 4 Markov Chains, with 5000 iterations for the first sample and 10 000 iterations for the second one. The first 50 per cent of each chain was discarded as "burn-in". This work is the first to perform a full HBM to study correlations in general and the correlation between the transit depth and the metallicity of *Kepler's* giant planets, in particular. Most importantly, the quantification of the intrinsic dispersion which has not been characterized before, is now defined.

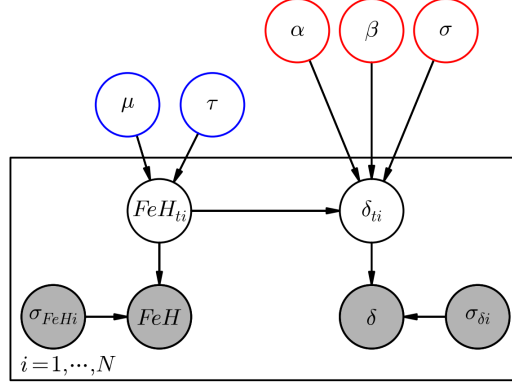


Figure 1. A graphical presentation of our HBM is given in this illustration. The gray nodes are the observed parameters. The true missing parameters are in the white nodes. The blue (upper left) nodes are the nuisance parameters and the red nodes (upper right) are the parameters of interest. FeH_i = stellar parameter of the i^{th} planet, $\sigma_{FeH,i}$ = uncertainty on the stellar metallicity of the i^{th} planet, δ_i = transit depth of the i^{th} planet, $\sigma_{\delta,i}$ = uncertainties on the transit depth of the i^{th} planet, FeH_{ti} = true stellar metallicity of the i^{th} planet, δ_{ti} = true transit depth of the i^{th} planet, μ and τ = nuisance parameters, α , β and σ = parameters of the linear model

4. Results

The posterior distributions for each of the parameters of interest (α , β and σ) produced by running MCMC are shown in the left panel in the Figure 2, for the two samples. The equation of the "best-fit" linear models are : $\delta = (0.07 \pm 0.014) + (0.02 \pm 0.14)FeH$ with an intrinsic scatter of $\sigma = 0.03 \pm 0.005$ for the confirmed planets sample. For the Complet subsample the best fit and the intrinsic scatter are $\delta = (0.13 \pm 0.06) + (-0.18 \pm 0.06)FeH$ and $\sigma = 0.05 \pm 0.02$, respectively. The transit depth is plotted against the metallicity of the host star along with their uncertainties, in the right panel of the Figure 2. The "best-fit" models (dotted lines) are shown along with 100 random samples from the MCMC chain. This shows clearly the existence of a large intrinsic scatter. A robust way to demonstrate the absence of correlation between the transit-depth and the stellar metallicity.

5. Discussion

We presented for the first time in the exoplanet literature, and within a Bayesian framework, a study of the correlation between the transit depth of *Kepler's* giant planets and the metallicity of the host star. Data from *Kepler*(Q1-Q17) ([7]) allowed us to characterize the intrinsic scatter in the relation with a robust statistical analysis. We did not assume that the observed parameters have negligible uncertainties. Moreover, we did not consider all the planets in the samples as *bona fide*. We expand the hierarchical model presentend in [9], to account for the false positive rates. We also considered in the model the relevant selection effects. The use of Markov Chain Monte Carlo (MCMC) to fit models to observations is becoming a standard practice in astronomy. We performed MCMC using the package PyStan to estimate the parameters of the hierarchical linear model. We established that there is no correlation between the transit depth and the stellar metallicity of *Kepler's* gas giant planets. Our model indicates that there is a relatively large intrinsic scatter in the relation. Hence, the previous results could probably be an artifact which shows the importance of accounting for uncertainties and for possible false positives. This is an exciting result actually. It also proves the importance of accounting for selection effects and biases within the transit surveys, such as *Kepler*, and the significance of

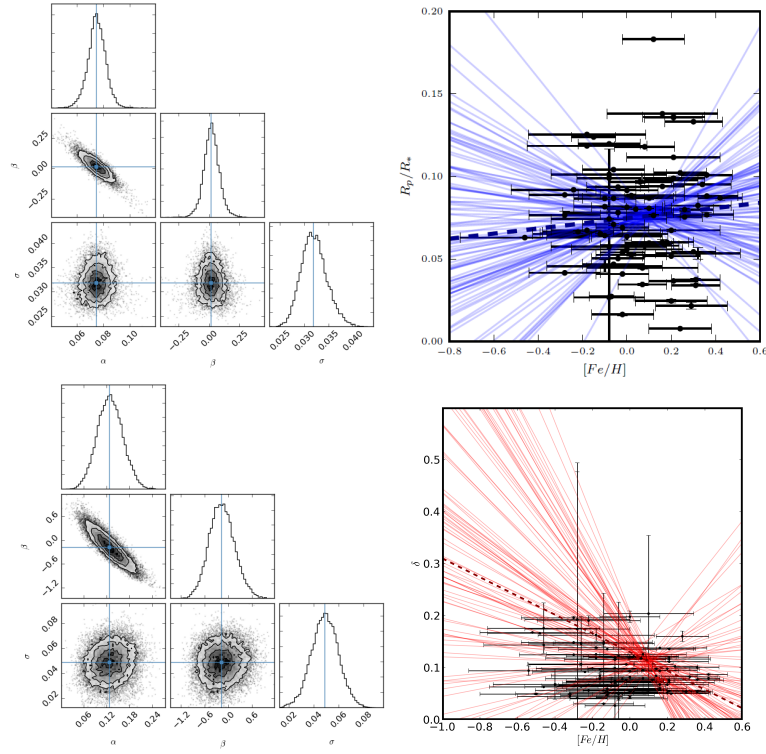


Figure 2. **Left** column are Posterior probability distribution for the parameters of our model and as computed by the MCMC package, PySTAN, marginalized over the other parameters (left up for the confirmed sample and left down for the complete one). **Right** column graphs are the transit depth (δ) of *Kepler*'s giant planets vs. the metallicity of the host star ($[Fe/H]$). The dashed lines represent the best fit line and the light lines are samples from the MCMC chain.

studying a complete subsample. [4] interpreted the negative correlation as evidence that metal-rich planets of a given mass are denser than their metal-poor counterparts, leading to smaller radii ([6]). On the contrary, with our robust statistical model, we have proven the independence of transit depth and the stellar metallicity. It certainly warrants further investigations to check what planetary formation model can explain the outcome.

References

- [1] Fischer D A and Valenti J **2005** *APJ* **622** 1102
- [2] Wang J and Fischer D A **2013** *arXiv:astro-ph/1310.7830*
- [3] Miller N and Fortney J J **2011** *APJ* **736** L29
- [4] Dodson-Robinson S E **2012** *APJ* **752** 72
- [5] Borucki W J, Koch D G, Basri G, Batalha N *et al* **2011** *APJ* **736** 19
- [6] Fortney J J and Nettelmann N **2010** *Space Sci. Rev.* **152** 423
- [7] Mullaly F, Coughlin J L, Thompson S E, Rowe J *et al* **2015** *APJS* **217** 31
- [8] Gaidos E and Mann A **2013** *APJ* **762** 81
- [9] Kelly B C **2007** *APJ* **665** 1489
- [10] Schaufman K C and Laughlin G **2011** *APJ* **738** 177
- [11] Huber D *et al* **2014** *APJS* **211** 2
- [12] Petigura E A, Howard A W and Marcy G W **2013** *PNAS* **110** 19273
- [13] Fressin F *et al* **2013** *APJ* **766** 81